Labor_03

LOGISTIC REGRESSION LINEAR CASE

Classification problem

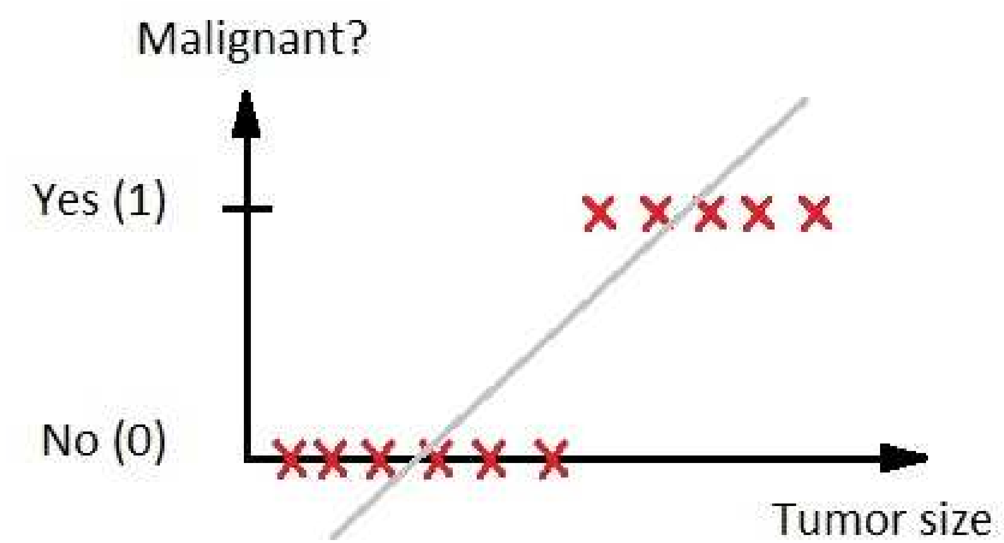
$y \in \{0, 1\}$ 0: Negative class
1: Positive class

$h_w(x) = XW$

Threshold classifier output $h_w(x)$ at 0.5:
If $h_w(x) \geq 0.5$, predict "y=1"
If $h_w(x) < 0.5$, predict "y=0"

$h_w(x)$ can be >1 or <0

Logistic regession: $0 \leq h_w(x) \leq 1$



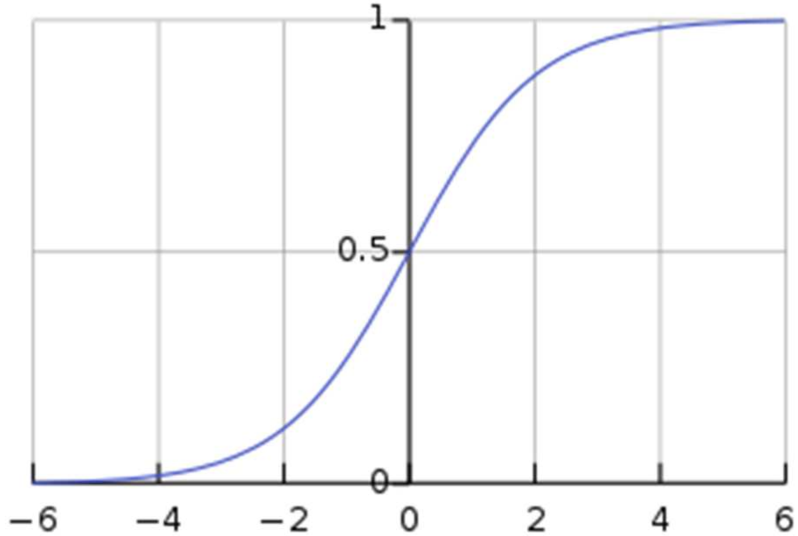
New Model

Logistic Regression Model:

Want $0 \leq h_w(x) \leq 1$

$h_w(x) = g(Xw)$

$$g(z) = \frac{1}{1 + e^{-z}}$$

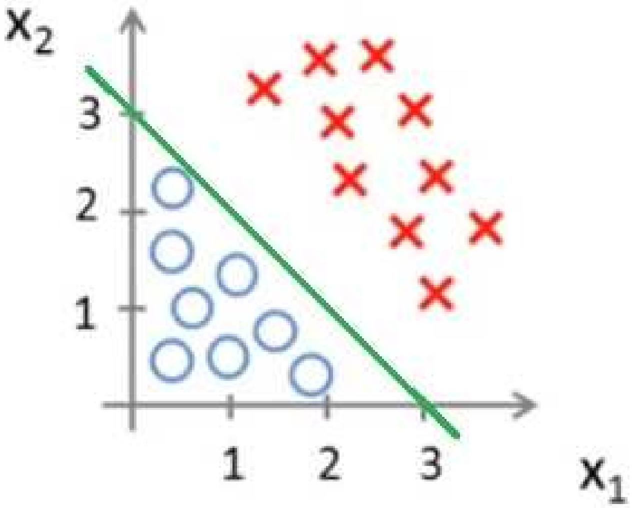


sigmoid function

$h_w(x) \Rightarrow$ estimated probability that $y = 1$ on input x

$h_w(x) = P(y = 1|x, W)$

Example linear boundary



$$h_w(x) = g(w_0 1 + w_1 x_1 + w_2 x_2)$$

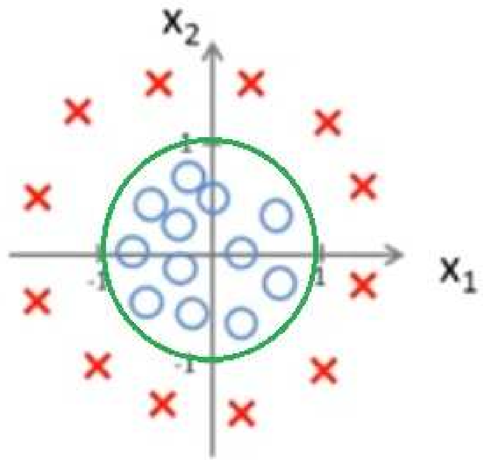
$$w = [-3\ 1\ 1]$$

$g(z) \geq 0.5$
when $z \geq 0$

Predict: $y = 1$ if $-3 + x_1 + x_2 \geq 0$

$$x_1 + x_2 \geq 3$$

Example non linear boundary



$$h_w(x) = g(w_0 1 + w_1 x_1 + w_2 x_2 + w_3 x_1^2 + w_4 x_2^2)$$

$$w = [-1\ 0\ 0\ 1\ 1]$$

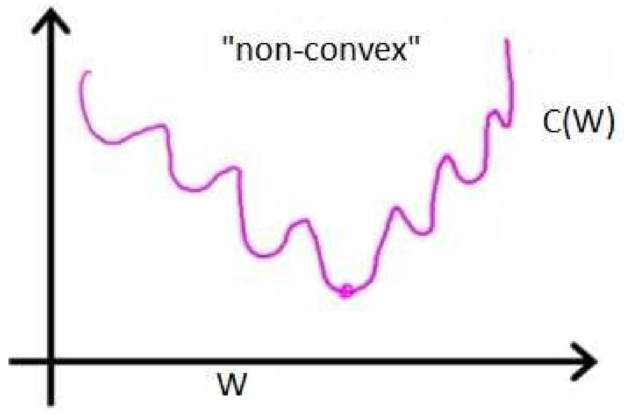
Predict: $y = 1$ if $-1 + x_1^2 + x_2^2 \geq 0$

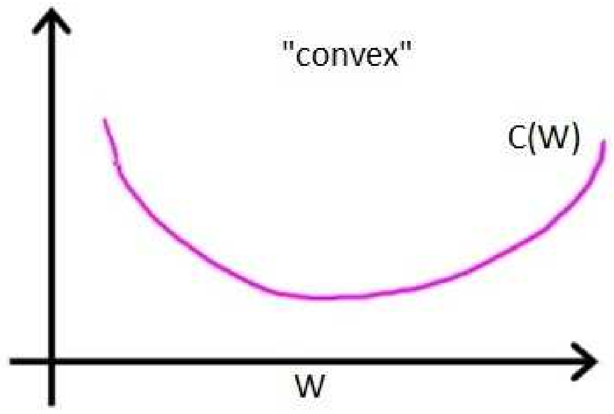
$$x_1^2 + x_2^2 \geq 1$$

Cost function

$$C = \frac{1}{2m}\sum_{i=1}^{m}(h_w(x^i) - y^i)^2$$

?



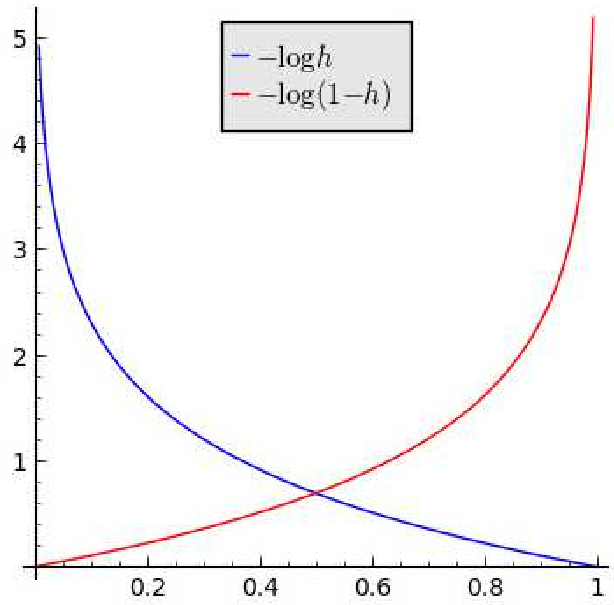


Cost function

$$Cost(h_w(x), y) = \begin{cases} -log(h_w(x)), \ if \ y = 1 \\ -log(1 - h_w(x)), \ if \ y = 0 \end{cases}$$

$$Cost(h_w(x), y) = -y \cdot log(h_w(x)) - (1 - y) \cdot log(1 - h_w(x))$$

$$C(w) = -\frac{1}{m} \sum_{i=1}^{m} y^i \cdot log(h_w(x^i)) - (1 - y^i) \cdot log(1 - h_w(x^i))$$



Gradient Descent

$$C(W) = -\frac{1}{m}\sum_{i=1}^{m} y^i \cdot log(h_w(x^i)) - (1-y^i)\cdot log(1-h_w(x^i))$$

Want min $min_W\{C(W)\}$
Algorithm:

repeat until convergence {

$$W_j := W_j - \mu \tfrac{\partial}{\partial W_j} C(W)$$

}

$$\frac{\partial}{\partial W_j} C(W) = \frac{1}{m}\sum_{i=1}^{m} (h_w(x^i) - y^i)\cdot x_j^i$$

Algorithm looks identical to linear regression!